

Deriving Leafless Trees and Urban Structures from the DLR 3K Airborne Camera System for the City of Braunschweig

Thomas Krauß

DLR, Remote Sensing Institute, Münchener Str. 20, 82234 Oberpfaffenhofen, Germany

ABSTRACT

In this work we present a method for detecting leafless trees in digital elevation models derived from high resolution aerial imagery. No spectral classification as usually conducted is needed for our approach. This makes the method applicable for stereo imagery using only panchromatic or RGB imagery or even acquisitions in early spring showing trees without foliage. The method is based on statistics of densely calculated aspect and slope angles in the DSM. Using a segmentation based on the watershed transformation gives areas large enough for reasonable statistics but also limited to individual tree or building objects. Removing the detected trees from the DSM allows the calculation of a building mask which can in turn be used for the urban modeling of buildings. The detected trees will be modeled in parallel and the models joined together afterwards. The development of the method and the evaluation is based on images acquired with the 3K camera system developed by DLR in early spring 2015 over the german city of Braunschweig. After description of the algorithm the selection of the main classification parameter is shown together with a manual quality assessment.

Keywords: 3K-Camera, RGB Data, Leafless Trees, Digital Elevation Models, Automatic Preprocessing, Urban Objects

1. INTRODUCTION

In this paper we present a method for detecting leafless trees based only on their structure in the digital surface model (DSM). The background of the study is the derivation of a urban city model from airborne imagery. The available dataset over the german city of Braunschweig contains only plain RGB images without a near infrared band (NIR) and were also acquired in early spring showing only trees without leaves.

For the detection of urban objects typically first a classification to high and low objects is done and the high objects are split to buildings and trees. This discrimination is usually done using a vegetation index like the NDVI (normalized difference vegetation index) based on the red and NIR bands of the image which shows clearly vivid vegetation. For the available dataset this approach is not possible due to (a) no available NIR band and (b) no leaves on the trees. Therefore also using other spectral indices like the GLI (green leaf index, see¹) fails due to the missing leaves.

Due to this restrictions a new approach for the discrimination of trees and buildings based only on structural characteristics of the DSM has to be developed for the subsequent task of urban model generation.

There exist many work on tree extraction from LiDAR data or multispectral imagery with or without DSMs.²⁻⁴ In most cases the imaged area contains only trees. Only few works address also the discrimination between buildings and trees in urban areas.⁵ Such approaches are usually based on the calculation of the NDVI (see 1) to divide trees and buildings in high elevated objects.

Further author information: (Send correspondence to Thomas Krauß)
E-mail: thomas.krauss@dlr.de

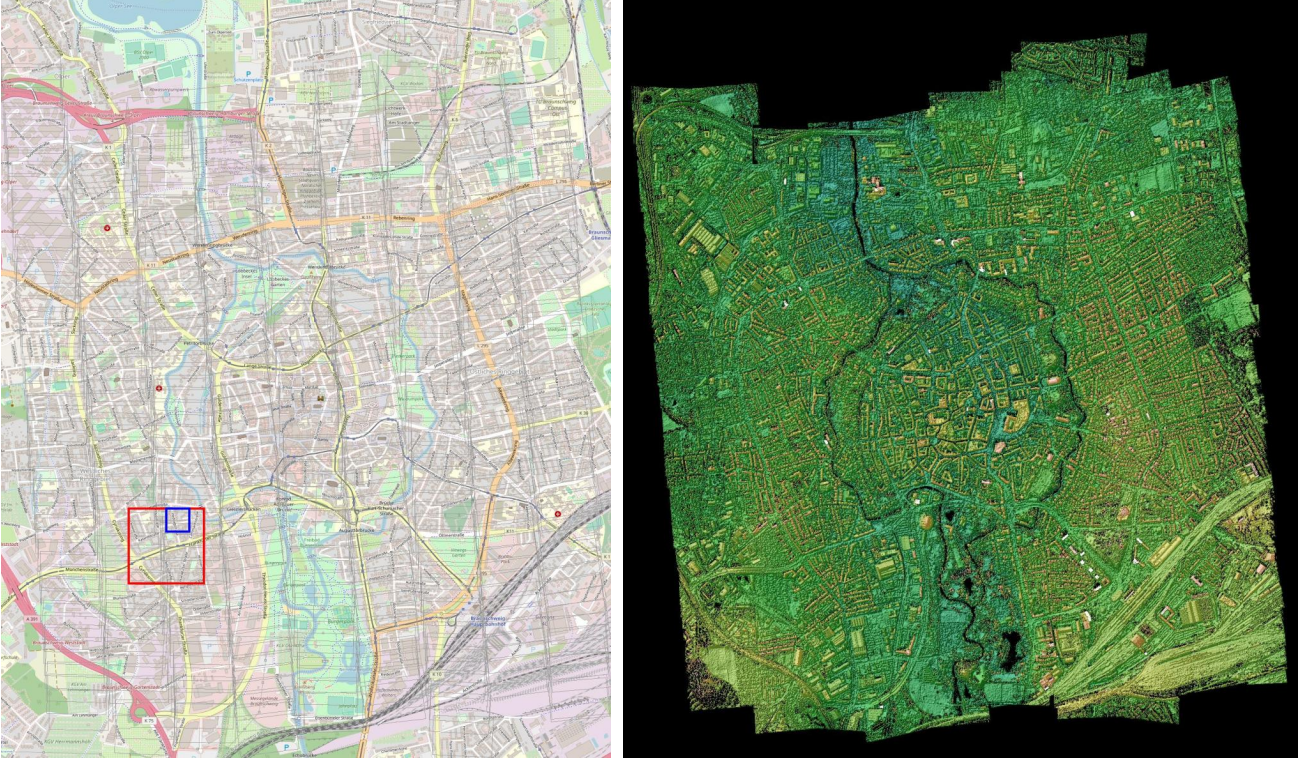


Figure 1: Overview of the acquired nadir images over Braunschweig with the DLR 3K camera on 2015-03-18 (left) and the derived DSM (right), $4.7 \times 5.3 \text{ km}^2$, GSD 0.1 m; The section used for the following investigations ($500 \times 500 \text{ m}^2$) is marked in red and in blue the detail area ($150 \times 150 \text{ m}^2$) with a road covered by trees

2. DATA

The imagery used for this investigation was acquired in a flight campaign over the german city of Braunschweig with the airborne 3K camera system on 2015-03-18 from 10:23 to 11:06 GMT. The system acquired 472 image triples – each left, nadir and right viewing. Fig. 1, left shows the footprints of the 472 nadir images over the area. Also shown in red is the section used for visualization in this paper together with a detail section in blue containing a road covered by trees joining the buildings on both sides of the road to one large elevated object.

The DLR developed 3K camera system is build up of three standard consumer Canon Mark III cameras delivering three 21 Mpx RGB images per second. The three cameras are mounted together with a GPS receiver and an IMU on a carrier in such a way one camera looks nadir, one to the left and one to the right. The cameras cover simultaneously an area up to 1.2 km wide and 300 m long with an overlap of about 70 % in flight direction when flying about 550 m above ground. The resulting images are pure three band (red, green and blue) uncalibrated JPG images. The cameras have no near infrared band (NIR) but moreover there were mostly no leaves on the trees due to the flight in early spring.

3. METHOD

Based on the images described in 2 an urban model should be derived. This includes the calculation of a digital surface model (DSM) from all the images after image bundle adjustment (see fig. 2, b) using the Semi-Global-Matching.⁶ Projecting the original RGB images on this DSM provide the true ortho mosaic as shown in fig. 2, a.

Based on the DSM a DTM (digital terrain model) representing the ground is derived (fig. 2, c) together with a so called normalized digital elevation model (nDEM) representing only the elevated objects (fig. 2, d). Based on the DSM, DTM and true ortho mosaic an urban model containing buildings, trees, water, roads and other objects should be derived.



Figure 2: True ortho image (a), calculated DSM (b), DTM (c) and nDEM containing only the elevated objects (d), of the section ($500 \times 500 \text{ m}^2$)

But as can be seen in fig. 2, (d) in the top right corner trees join buildings across a street to one big elevated object, which cannot be splitted into separated objects. Normally this is done using the so called NDVI – the normalized difference vegetation index – which is calculated from the red (R_r) and NIR (R_n) bands of the image as

$$NDVI = \frac{R_n - R_r}{R_n + R_r}. \quad (1)$$

Due to the missing NIR band in the data this approach cannot be followed. For three band RGB imagery there is an other approach¹ – the green leaf index (GLI) which is defined as an index involving the blue (R_b), green (R_g) and red (R_r) bands as

$$GLI = \frac{2 \cdot R_g - R_r - R_b}{2 \cdot R_g + R_r + R_b}. \quad (2)$$

But in our case for the March flight campaign this approach fails also due to the missing leaves of the trees as can be seen in fig. 3.

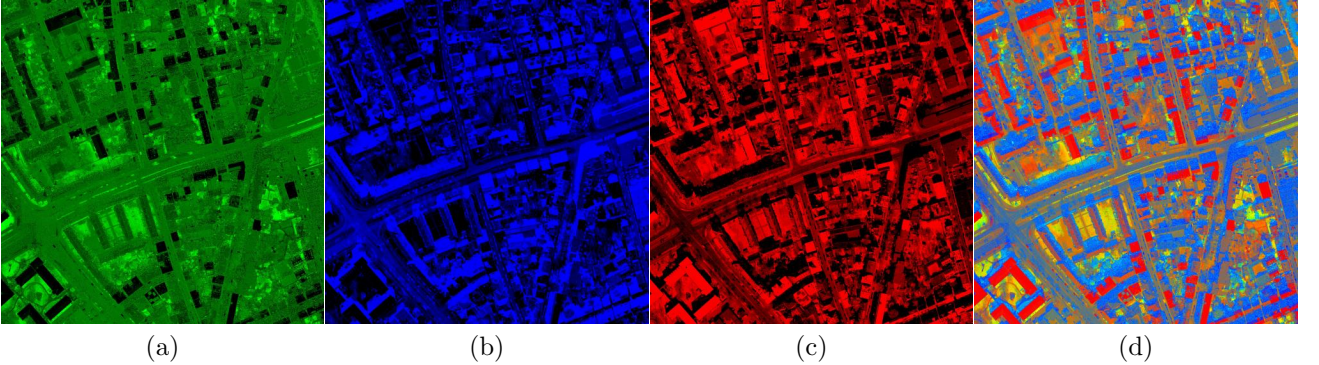


Figure 3: Computed GLI of the test area (a, green), shadow (b, blue), sealed (c, red) and combined masks (d)

Fig. 3 shows the calculated GLI in green (a). Also analogous a shadow mask SHD (b) and a sealed mask SM (c) were calculated as

$$SHD = \frac{1}{3} \cdot \left(\frac{R_b - R_g}{R_b + R_g} + 2 \cdot \frac{R_b - R_r}{R_b + R_r} \right) \quad \text{and} \quad SM = \frac{1}{3} \cdot \left(\frac{R_r - R_g}{R_r + R_g} + 2 \cdot \frac{R_r - R_b}{R_r + R_b} \right) \quad (3)$$

Combining all three bands give the class image in fig. 3, right. Yellow areas depict “stressed” vegetation. As can be seen in the top right detail area the trees in the road show no green (GLI) properties but look rather “sealed”.

So a new method for detecting leafless trees from the DSM was needed. The developed method presented in this paper is based only on structural properties of the DSM. In a first step for each point of the DSM the aspect and slope angles are calculated (see fig. 4). The aspect angle gives the direction downward and runs from 0° (north) over 90° (east), 180° (south) and 270° (west) back to 360° (north). The slope runs from 0° (flat) to 90° (vertically steep).

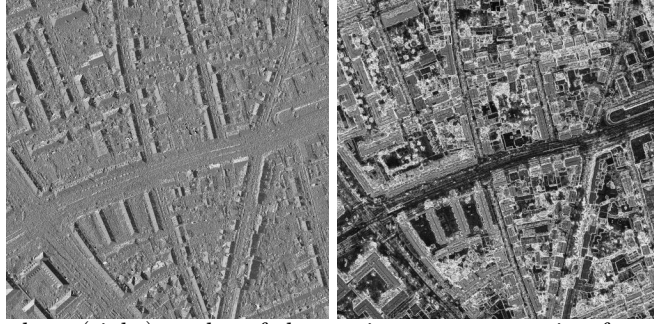


Figure 4: Aspect (left) and slope (right) angles of the section, aspect running from 0° (black, north) over east, south, west to 360° (white) and slope running from 0° (black, flat) to 90° (white, steep)

Due to the especially in shaded areas noisy DSM the directions are partly also very noisy. To do a valid analysis some statistical approaches have to be applied. This is done first by filtering the DSM using a gaussian filter. The later analysis is based mainly on the filter size σ of this filter as parameter running from 10 cm (one pixel) up to 11.5 m. Fig. 5 shows the original filled DSM (a) together with the DSM Gauss-filtered using $\sigma = 1.1 \text{ m} = 11 \text{ px}$ (b).

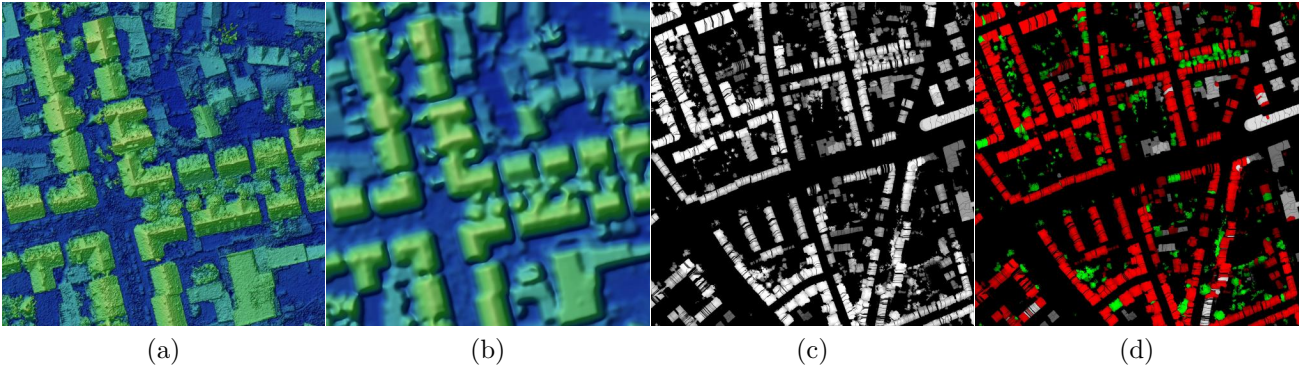


Figure 5: DSM (a) and gaussian filtered ($\sigma = 1.1 \text{ m}$) DSM (b), watershed regions (c), classification of watershed regions (d, roof: red, tree: green, flat: gray)

For statistical analysis larger areas representing ideally single objects are needed. Since both buildings as well as trees show local maxima for each object in the gaussian filtered DSMs a watershed-transformation⁷ as commonly used for tree-detection can be applied. As this operation fills up sinks up to watersheds between we need to invert our DSM to make rooftops and treetops to local minima. Fig. 5, c shows the segmentation based on the watershed-transformation limited to high objects.

For each of the segments statistics for all aspect and slope angles of the segment are calculated. Fig. 6 shows typically resulting histograms of the angular statistics. Shown are 2D-histograms with the aspect angle modulo 180° in horizontal direction (left is 0 or 180° running to right with 180 or 360°). The slope angle runs in vertical direction from top (0°, flat) to bottom (90°, steep). Black means no pixels of this angular aspect-slope-

combination, white means maximum number in this aspect-slope-bin. For each direction 30 bins are used – i.e. 6° per bin in aspect- and 3° in slope-direction.

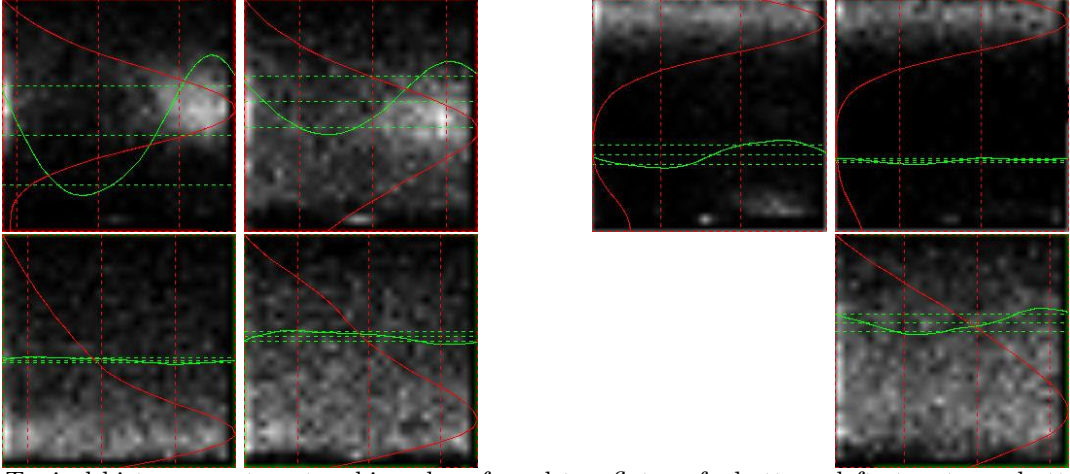


Figure 6: Typical histograms; top: two hipped roofs and two flat roofs; bottom, left: two trees; bottom, right: a roof misclassified as tree; the histograms show the azimuth angle modulo 180° from left to right and the slope angle from top (0° , flat) to bottom (90° , steep), green: histogram of aspects, red: histogram of slopes, dotted lines: mean and mean $\pm \sigma$

Overlayed over the 2D-histograms in fig. 6 are single histograms in aspect (green) and slope-direction (red) together with the mean and mean plus/minus standard deviations of the histogram counts as dotted lines.

Based on these histograms we try to classify each of the watershed-segments into one of three classes:

- Roof: a slanted roof is characterized by sharp maxima in aspect as well as in slope direction
- Flat: a flat roof is characterized by a random distribution in aspect angles but a maximum of slopes near zero
- Tree: a (leafless) tree is characterized by a (nearly) random distribution in aspect angles but a maximum of steep slopes due to single branches in different heights

Based on these criteria each watershed segment is assigned a class roof (red), tree (green) or flat (gray) as shown in fig. 5 (d).

As can be seen in fig. 6, right-bottom, some roof segments show no clear statistics due to too many errors and noise in the DSM generation process. Such areas will be mis-classified as trees.

4. RESULTS

Fig. 7 shows the classification results for experiments conducted with sizes of the gaussian filter ranging from 0.3 m up to 7.9 m. As can be seen too small σ result in many small segments with partly bad statistics. On the other hand a too large filter size merges too much segments together. The optimum filter size for a good statistics in large enough segments but not joining too much different objects was found for $\sigma = 1.1 \text{ m} = 11 \text{ px}$.

Table 1 shows the analysis for the detail area shown in fig. 7 and $\sigma = 1.1 \text{ m}$. The watershed transformation splitted the area up to 206 elevated segments. Automatic classification resulted in 52 trees, 121 roofs and 33 flat roofs. Manually verification shows there were 53 trees, 118 roofs and 35 flat roofs. Table 1 shows that there were mis-classifications mostly between roofs and trees but also to a minor degree with flat objects.

In summary the completeness of detected trees is $47/53 = 88.7 \%$, the correctness $47/52 = 90.4 \%$. This means for our use-case (removing tree objects for urban modelling) we miss 5 of 118 buildings (4.2 %) but have still 6 of 53 trees (11.3 %) classified wrongly as buildings. Over all $17/206 = 8.25 \%$ of all objects will be wrong.

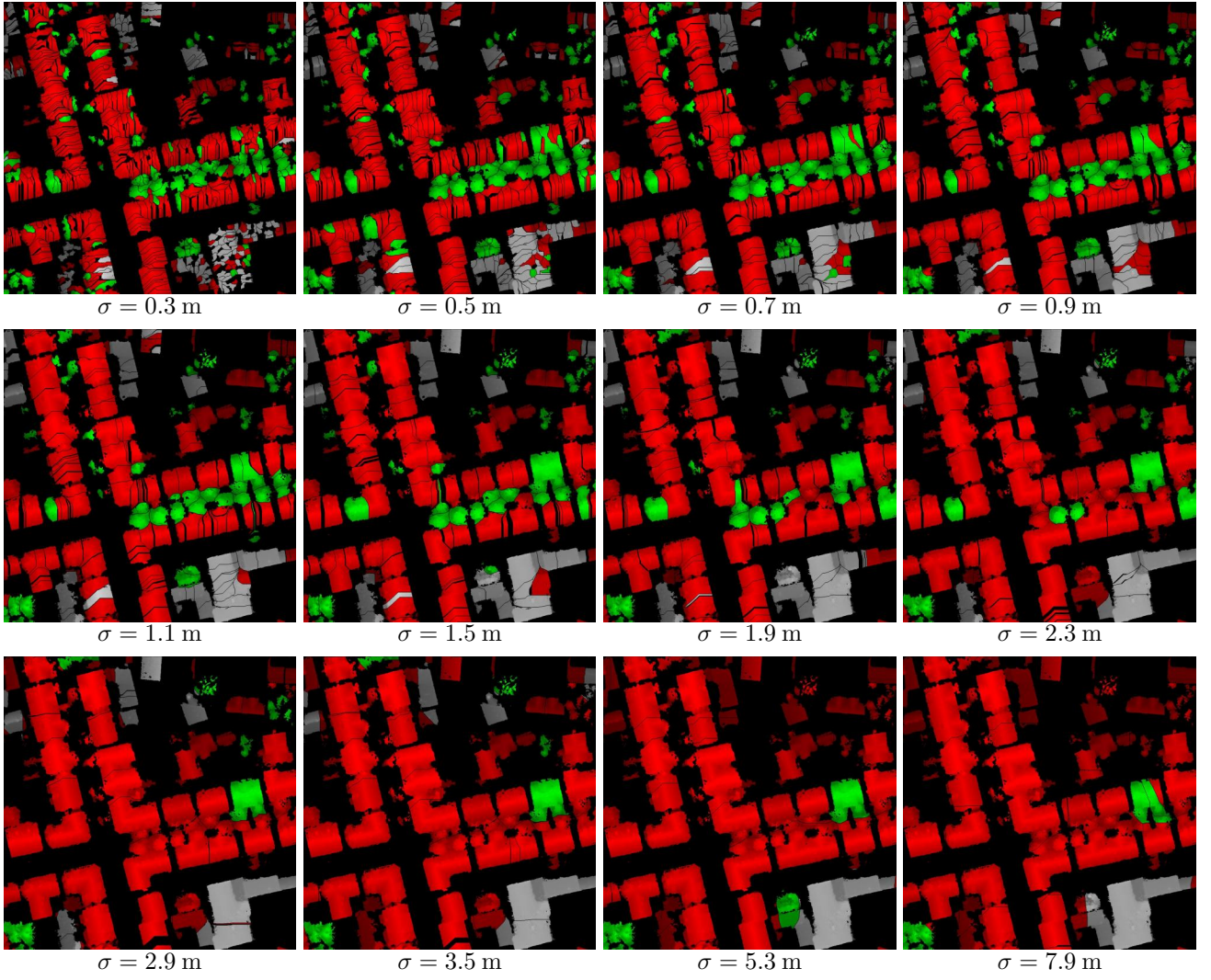


Figure 7: Results of the detail section $150 \times 150 \text{ m}^2$ based on Gaussian filtered DSMs with given σ ; classes: red: roof, green: tree, gray: flat, brightness represents the height of the nDEM

Table 1: Exemplary results for detail section using $\sigma = 1.1 \text{ m}$

| Ground-Truth | Detected | | | |
|--------------|----------|------|------|-----|
| | tree | roof | flat | |
| tree | 47 | 4 | 2 | 53 |
| roof | 5 | 112 | 1 | 118 |
| flat | 0 | 5 | 30 | 35 |
| | 52 | 121 | 33 | |

Fig. 8 shows a detail of a region containing different types of trees. So all objects are really trees as can be verified with the associated ortho image. From the ten trees in this area six were correctly classified as trees, two mis-classified as roofs and even one as flat roof.

Finally fig. 9 shows the resulting classification using the optimal $\sigma = 1.1 \text{ m}$ for the whole section (left) and the derived building mask (right). For generation of the building mask first all detected tree watershed segments are removed from the nDEM followed by a morphologically gray value opening and closing to clean up small artefacts. Fig. 10 shows the resulting derived vector objects of the urban objects (here only buildings and trees)

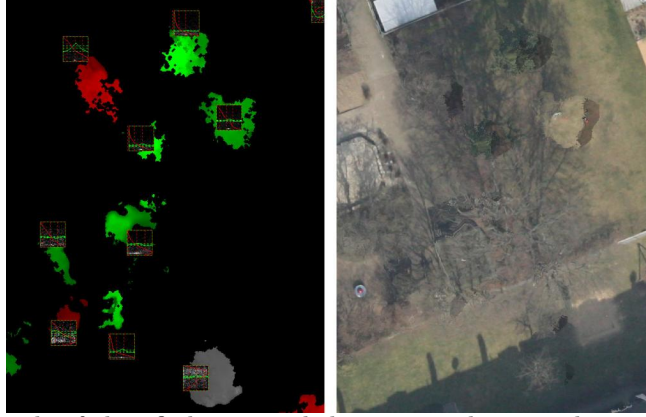


Figure 8: Example of classified trees and the associated true ortho image as reference

for the input data without the developed tree detection and building mask generation shown in this paper (left) and with removed trees (right) which were re-inserted as vector objects after the building detection.

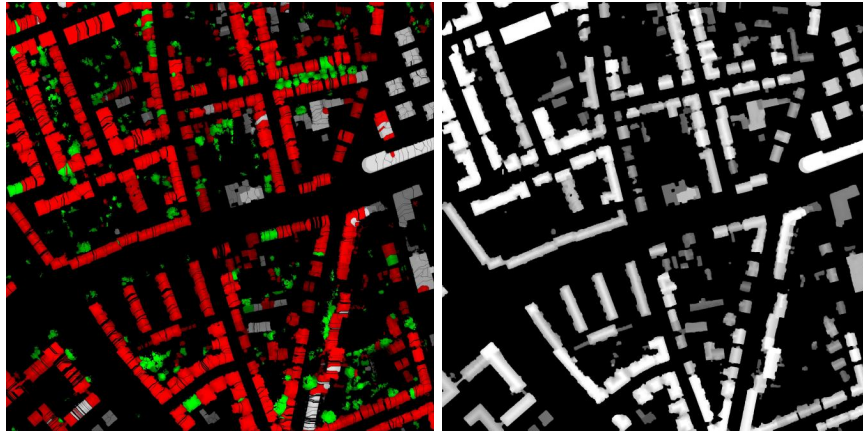


Figure 9: Results for $\sigma = 1.1$ m; left: classification (red: roof, green: tree, gray: flat), right: morphologically filtered building mask with tree removed

5. CONCLUSION AND OUTLOOK

In this work we presented a newly method for detecting trees only based on their structural properties – especially the statistics of aspect and slope angles – in the digital surface model (DSM). The method needs as main parameter the filter size of a gaussian filter for smoothing the DSM to eliminate fine structures of trees transforming them to a smooth dome like shape.

Deeper analysis show that a filter size of $\sigma = 1.1$ m yields the best results for the available airborne imagery with a resolution of 0.1 m. Still there are objects which get mis-classified. Mostly these are buildings with a very noisy roof surface due to bad illumination conditions or other errors in the generation of the dense DSM. Such buildings get classified wrongly as trees.

An other problem are lower trees beneath roofs which get joined together to one object by the watershed transformation. Such segments get mostly classified as trees. Some few trees were classified as roofs, one even as a flat roof.

Over all the completeness of detected trees is 88.7 % with a correctness of 90.4 %. For the urban modelling 4.2 % of buildings get wrongly removed by the method were 11.3 % of the trees were not removed and get modeled as buildings. All together about 8.25 % of all modeled objects will be wrong.

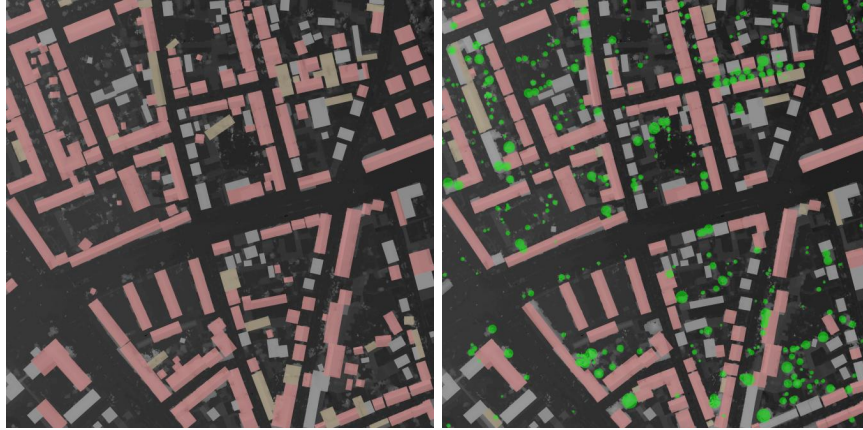


Figure 10: Results for $\sigma = 1.1$ m; left: originally wrongly derived objects without removing trees, right: detected objects (buildings and trees) after applying the tree detection presented in this paper

Using more sophisticated segmentation than only the watershed transformation on the smoothed DSM may result in better object delineation and clearer statistics. Also using additionally knowledge of neighbouring segments and joining connected segments while separating trees in a more sophisticated way from roof segments may result in better classification rates.

Also the possibility of applying the method to space-borne DSMs may be investigated in future.

REFERENCES

- [1] Hunt, E. R., Doraiswamy, P. C., McMurtrey, J. E., Daughtry, C. S. T., Perry, E. M., and Akhmedov, B., “A visible band index for remote sensing leaf chlorophyll content at the canopy scale,” *International Journal of Applied Earth Observation and Geoinformation* **2013**(21), 103–112 (2013).
- [2] Pollock, R. J., *The Automatic recognition of Individual trees in Aerial Images of Forests Based on a Synthetic Tree Crown Image Model*, PhD thesis, The University of British Columbia (6 1996).
- [3] Wang, L., Gong, P., and Biging, G. S., “Individual Tree-Crown Delineation and Treetop Detection in High-Spatial-Resolution Aerial Imagery,” *Photogrammetric Engineering & Remote Sensing* **70**(3), 351–357 (2004).
- [4] Straub, B.-M., *Automatische Extraktion von Bäumen aus Fernerkundungsdaten*, PhD thesis, Universität Hannover (2003).
- [5] Straub, B. M., “Automatic extraction of Trees from aerial images and surface models,” in [*ISPRS Archives*], *ISPRS Archives* **XXXIV**, 157–164 (9 2003).
- [6] Hirschmüller, H., “Accurate and efficient stereo processing by semi-global matching and mutual information,” in [*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], *Computer Vision and Pattern Recognition* **2**, 807–814 (2005).
- [7] Beucher, S., “Watersheds of Functions and Picture Segmentation,” in [*IEEE International Conference on Acoustics, Speech and Signal Processing*], *IEEE International Conference on Acoustics, Speech and Signal Processing* , 1928–1931 (1982).